LIST OF THE CLAIMS

1. (Currently Amended) A method for visual-based recognition of an object, said method comprising:

receiving digital depth data for at least a pixel of an image of an object, which is not required to be inside of a subject, said depth data comprising information relating to a distance from a visual sensor to a portion of said object shown at said pixel, said visual sensor comprising an emitter and sensor of light, wherein said light is selected from the group of electromagnetic radiation consisting of visible light, infrared light, and ultraviolet light and wherein said receiving of said depth data does not require special behavior from one of said object and said subject;

generating a two dimensional plan-view image based in part on said depth data, wherein said generating includes generating said plan-view image as if said object [[were]] was viewed at an axis normal to ground level from above and wherein generating other view images based on different orientations of said object other than at said axis normal to ground level from above is not required;

extracting a plan-view template from said plan-view image, wherein at least a portion of said plan-view image is transformed; and

processing said plan-view template at a classifier, that is executing on a computer system, to assign a class to said plan-view template, wherein said classifier is trained to make a decision according to pre-configured parameters determined at least in part based on said class of said plan-view template.

2. (Original) The method as recited in Claim 1 further comprising receiving non-depth data for said pixel.

3. (Original) The method as recited in Claim 1 wherein said visual sensor determines said depth data using stereopsis based on image correspondences.

4. (Original) The method as recited in Claim 1 wherein said generating said plan-view image comprises selecting a subset of said depth data based on foreground segmentation.

5. (Previously Presented) The method as recited in Claim 1 wherein said generating said plan-view image further comprises:

generating a three-dimensional point cloud of a subset of pixels based on said depth data, wherein a point of said three-dimensional point cloud comprises a three-dimensional coordinate;

partitioning said three-dimensional point cloud into a plurality of vertically oriented bins; and

mapping at least a portion of points of said plurality of vertically oriented bins into at least one said plan-view image based on said three-dimensional coordinates, wherein said plan-view image is a two-dimensional representation of said three-dimensional point cloud comprising at least one pixel corresponding to at least one vertically oriented bin of said plurality of vertically oriented bins.

6. (Original) The method as recited in Claim 4 further comprising receiving non-depth data for said pixel, and wherein said foreground segmentation is based at least in part on said non-depth data.

7. (Original) The method as recited in Claim 5 further comprising dividing said three-dimensional point cloud into a plurality of slices, and wherein said generating said plan-view image is performed for at least one slice of said plurality of slices.

8. (Previously Presented) The method as recited in Claim 7 wherein said extracting a plan-view template from said plan-view image further comprises extracting a plan view template from at least two plan-view images corresponding to different slices of said plurality of slices, wherein said plan-view template comprises a transformation of at least said portion of said plan-view images, such that said plan-view template is processed at said classifier.

9. (Original) The method as recited in Claim 1 wherein said extracting said plan-view template from said plan-view image is based at least in part on object tracking.

10. (Original) The method as recited in Claim 1 wherein said classifier is a support vector machine.

11. (Original) The method as recited in Claim 2 wherein said plan-view image is based in part on said non-depth data.

12. (Original) The method as recited in Claim 1 wherein said object is a person.

13. (Original) The method as recited in Claim 1 wherein said plan-view image comprises a value based at least in part on an estimate of height of a portion of said object above a surface.

14. (Original) The method as recited in Claim 1 wherein said plan-view image comprises a value based at least in part on color data for a portion of said object.

15. (Original) The method as recited in Claim 1 wherein said plan-view image comprises a value based at least in part on a count of pixels obtained by said visual sensor and associated with said object.

16. (Original) The method as recited in Claim 1 wherein said plan-view template is represented in terms of a vector basis.

17. (Original) The method as recited in Claim 16 wherein said vector basis is obtained through principal component analysis (PCA).

18. (Original) The method as recited in Claim 13 further comprising performing height normalization on said plan-view template.

19. (Original) The method as recited in Claim 1 wherein said decision is to distinguish between a human and a non-human.

20. (Original) The method as recited in Claim 1 wherein said decision is to distinguish between a plurality of different human body orientations.

21. (Original) The method as recited in Claim 1 wherein said decision is to distinguish between a plurality of different human body poses.

22. (Original) The method as recited in Claim 1 wherein said decision is to distinguish between a plurality of different classes of people.

23. (Currently Amended) A visual-based recognition system comprising:

a visual sensor for capturing depth data for at least a pixel of an image of an object, which is not required to be inside of a subject, said depth data comprising information relating to a distance from said visual sensor to a portion of said object visible at said pixel, said visual sensor comprising an emitter and sensor of light, wherein said light is selected from the group of electromagnetic radiation consisting of visible light, infrared light, and ultraviolet light and wherein said capturing of said depth data does not require special behavior from one of said object and said subject;

a plan-view image generator for generating a two dimensional plan-view image based on said depth data, wherein said generating of said plan-view image includes generating said plan-view image as if said object [[were]] was viewed at an axis normal to ground level from above and wherein generating other view images based on different orientations of said object other than at said axis normal to ground level from above is not required;

a plan-view template generator for generating a plan-view template based on said plan-view image; and

a classifier for making a decision concerning recognition of said object, wherein said classifier is trained to make said decision according to pre-configured parameters that were determined at least in part based on a class assigned to said plan-view template.

24. (Original) The visual-based recognition system as recited in Claim 23 wherein said visual sensor is also for capturing non-depth data.

25. (Original) The visual-based recognition system as recited in Claim 23 wherein said visual sensor determines said depth data using stereopsis based on image correspondences.

26. (Original) The visual-based recognition system as recited in Claim 23 wherein said plan-view image generator comprises a pixel subset selector for selecting a subset of pixels of said image, wherein said pixel subset selector is operable to select said subset of pixels based on foreground segmentation.

27. (Original) The visual-based recognition system as recited in Claim 23 wherein said classifier is a support vector machine.

28. (Original) The visual-based recognition system as recited in Claim 24 wherein said plan-view image is based in part on said non-depth data.

29. (Original) The visual-based recognition system as recited in Claim 23 wherein said plan-view image generator is operable to generate a three-dimensional point cloud based on said depth data, wherein a point of said three-dimensional point cloud comprises a three-dimensional coordinate.

30. (Original) The visual-based recognition system as recited in Claim 29 wherein said plan-view image generator is operable to divide said three-dimensional

point cloud into a plurality of slices such that a plan-view image may be generated for at least one slice of said plurality of slices.

31. (Previously Presented) The visual-based recognition system as recited in Claim 30 wherein said plan-view template generator is operable to extract a plan-view template from at least two plan-view images corresponding to different slices of said plurality of slices, wherein said plan-view template comprises a transformation of at least said portion of said plan-view images, such that said plan-view template is processed at said classifier.

32. (Currently Amended) A method for visual-based recognition of an object representative in an image, said method comprising:

generating a three-dimensional point cloud based on digital depth data for at least a pixel of an image of said object, which is not required to be inside of a subject, said depth data comprising information relating to a distance from a visual sensor to a portion of said object visible at said pixel, said visual sensor comprising an emitter and sensor of light, wherein said light is selected from the group of electromagnetic radiation consisting of visible light, infrared light, and ultraviolet light, said three-dimensional point cloud representing a foreground surface visible to said visual sensor and wherein a pixel of said three-dimensional point cloud comprises a three-dimensional coordinate and wherein said generating of said three-dimensional point cloud does not require special behavior from one of said object and said subject;

partitioning said three-dimensional point cloud into a plurality of vertically oriented bins;

mapping at least a portion of points of said vertically oriented bins into at least one said plan-view image based on said three-dimensional coordinates, wherein said plan-view image is a two-dimensional representation of said three-dimensional point cloud comprising at least one pixel corresponding to at least one vertically oriented bin of said plurality of vertically oriented bins, wherein said mapping includes generating said plan-view image as if said object [[were]] <u>was</u> viewed <u>at an axis normal to ground level</u> from above; and

processing said plan-view image at a classifier, that is executing on a computer system, wherein said classifier is trained to make a decision according to pre-configured parameters and wherein said pre-configured parameters were determined based at least in part on a class assigned to a plan-view template that was extracted from said plan-view image by transforming at least a portion of said plan-view image, said classifier does not require other view images based on different orientations than at said axis normal to ground level from above of said object in order to make said decision.

33. (Original) The method as recited in Claim 32 wherein said three-dimensional point cloud and said plan-view image are also based at least in part on non-depth data.

34. (Original) The method as recited in Claim 32 wherein said visual sensor determines said depth data using stereopsis based on image correspondences.

35. (Previously Presented) The method as recited in Claim 32 wherein said plan view template comprises a transformation of at least said portion of said plan view image, and such that said plan-view template is processed at said classifier.

36. (Original) The method as recited in Claim 32 further comprising dividing said three-dimensional point cloud of into a plurality of slices, and wherein said mapping at least a portion of points comprises mapping points within a slice of said plurality of slices of said three-dimensional point cloud into said plan-view image.

37. (Previously Presented) The method as recited in Claim 36 wherein said plan view template comprises a transformation of at least said portion of said plan view image, such that said plan-view template is processed at said classifier.

38. (Original) The method as recited in Claim 32 wherein said classifier is a support vector machine.

39. (Original) The method as recited in Claim 32 wherein said plan-view image is generated from a subset of pixels of said image selected based on foreground segmentation.

40. (Previously Presented) The method as recited in Claim 36 further comprising extracting a plan view template from at least two plan view images corresponding to different slices of said plurality of slices, wherein said plan view template comprises a transformation of at least said portion of said plan view images, such that said plan-view template is processed at said classifier.